ELSEVIER

Gene 193 (1997) 229–237

GENE
AN INTERNATIONAL JOURNAL ON GENES AND GENOMES

High-level production and purification of biologically active proteins from bacterial and mammalian cells using the tandem pGFLEX expression system

Herbert T. Manoharan, Jean Gallo, Andrew M. Gulick, William E. Fahl *

McArdle Laboratory for Cancer Research, University of Wisconsin, 1400 University Avenue, Madison, WI 53706, USA

Received 5 November 1996; received in revised form 16 January 1997; accepted 23 January 1997; Received by J. Wild

Abstract

Because of the complexities involved in the regulation of gene expression in *Escherichia coli* and mammalian cells, it is considered general practice to use different vectors for heterologous expression of recombinant proteins in these host systems. However, we have developed and report a shuttle vector system, pGFLEX, that provides high-level expression of recombinant glutathione *S*-transferase (GST) fusion proteins in *E. coli* and mammalian cells. pGFLEX contains the cytomegaloma virus (CMV) immediate-early promoter in tandem with the *E. coli lacZpo* system. The sequences involved in gene expression have been appropriately modified to enable high-level production of fusion proteins in either cell type. The pGFLEX expression system allows production of target proteins fused to either the N or C terminus of the GST π protein and provides rapid purification of target proteins as either GST fusions or native proteins after cleavage with thrombin. The utility of this vector in identifying and purifying a component of a multi-protein complex is demonstrated with cyclin A. The pGFLEX expression system provides a singular and widely applicable tool for laboratory or industrial production of biologically active recombinant proteins in *E. coli* and mammalian cells. © 1997 Elsevier Science B.V.

Keywords: Cyclin; Fusion protein; Glutathione *S*-transferase

1. Introduction

A wide variety of gene expression vectors are available for the production of heterologous proteins, enzymes, hormones, immunogenic molecules, antibodies and low molecular weight peptides in host systems as diverse as bacterial and mammalian cells. Several inducible expression systems exist, utilizing a variety of different promoters. Examples include the *Pac*, *Ptac*, *Ptrp*, *Plpp*, PT7 and λp_L for inducible bacterial expression, and the mouse mammary tumor virus LTR, or metallothionein promoter for inducible eukaryotic expression. The protein of interest can be expressed in native forms directly from its gene or cDNA, or produced as a fusion protein (see Gold, 1990). Fusion systems commonly used today include schistosomal GST (Smith and Johnson, 1988), maltose-binding protein, thioredoxin, Staphylococcal protein A, hemagglutinin tag and histidine tag (see Olins and Lee, 1993). The advantage of the GST fusion system is that the purification conditions are relatively mild compared to other systems. Prokaryotic expression systems have been engineered for efficiency and convenience; however, they have a number of limitations for the synthesis of eukaryotic proteins (Datur et al., 1993). Many eukaryotic proteins have been found to be difficult to produce or to lack biological activity when produced in *Escherichia coli*. Such problems may be due to mRNA or protein instability (Maurizi, 1992), improper protein folding, and formation of inclusion bodies (Mitraki et al., 1991). Often, these proteins pose fewer problems

* Corresponding author. Tel. +1 608 2627499; fax +1 608 2622824; e-mail: fahl@oncology.wisc.edu

Abbreviations: aa, amino acid(s); Ab, antibody(ies); Ap, ampicillin; βGal, β-galactosidase; BGH, bovine growth hormone; bp, base pair(s); cDNA, DNA complementary to RNA; GSH, glutathione; GST, glutathione *S*-transferase; *GST*, glutathione *S*-transferase encoding gene; HSV, herpes simplex virus; IPTG, isopropyl β-D-thiogalactopyranoside; *lacZpo*, *lac* promoter-operator; LB, Luria broth (medium); *LTR*, long terminal repeat; Luc, luciferase; MCS, multiple cloning site(s); Nm, neomycin; nt, nucleotide(s); oligo, oligodeoxyribonucleotide; ORF, open reading frame; PAGE, polyacrylamide-gel electrophoresis; ^{R}, resistant/resistance; RBS, ribosome-binding site; RSV, Rous sarcoma virus; SD, Shine-Dalgarno (sequence); SDS, sodium dodecyl sulfate; SV40, Simian virus 40; *tk*, thymidine kinase-encoding gene.

0378-1119/97/$17.00 © 1997 Elsevier Science B.V. All rights reserved.
PII S0378-1119(97)00126-1

when produced in eukaryotic cells. Thus, for many gene expression applications, the ability to synthesize a protein/fusion protein in both *E. coli* and mammalian cells would be quite useful. We have developed the pGFLEX shuttle vector system (Fig. 1) and demonstrate the utility of the system by synthesizing human cyclin A (Wang et al., 1990) and the firefly luciferase protein (de Wet et al., 1987) in both mammalian cells and *E. coli*.

2. Materials and methods

2.1. Construction of plasmids

The plasmids described herein were constructed in several subcloning steps as outlined in Fig. 1. Plasmid pGPi-2 (Kano et al., 1987), containing the entire coding region of the human *GST* π cDNA within *Eco*RI sites, was obtained from Dr. Masami Muramatsu. To introduce MCS on either side of the *GST* π cDNA, pGPi-2 was digested with *Eco*RI and the GST-π-cDNA was inserted into the *Eco*RI site of pIC-20H (Marsh et al., 1984). The resulting vector was designated as pIC-20H-GSTπ and the *GST* π cDNA was in reverse transcriptional orientation relative to the *lacZpo*. The *GST* π cDNA was brought into the proper transcriptional orientation and translational reading frame by excising the *GST* π cDNA from pIC-20H-GSTπ with *Cla*I-*Pst*I and inserting it into the *Cla*I + *Pst*I sites of pIC-20R. This insertion enabled the *GST* π cDNA to be expressed from the *Plac* resulting in a fusion protein with the βGal leader sequence at its N terminus. Conceptual translation of the nt sequence at the *Cla*I junction region of the *plac-GST* π cDNA was checked by nt and peptide sequence analysis (Fig. 1B). The N-terminal aa sequence of the βGal-GST π fusion protein was done with purified GST π using an Applied Biosystems model 470 protein sequenator (Hewick et al., 1981). In order to add the thrombin (T) cleavage and MCS sequences at the C terminus of GST π, a 769-bp DNA fragment containing the *Plac*, βGal leader sequence, and *GST* π coding region without the start codon (209 aa; 24.8 kDa) was amplified by PCR from pIC-20R-GSTπ using 5′-*Pvu*II (5′-GAGCTGCAGC TGGCACGACAGGTT) and 3′-thrombin-*Bam*HI (5′-TTCCGGGGATCCACGCGGA-ACCAGCTGTTTCCGTTGCCATTTGAT) primers and inserted between *Bsp*MI and *Bam*HI sites of pGEX-4T-1 backbone. The pGEX-4T-1 backbone was prepared by removing the fragment containing the *ptac* and schistosomal GST cDNA through digestion with *Bsp*MI + *Bam*HI. A 790-bp (*Pvu*II-*Apa*I) fragment containing the *Plac*, βGal leader sequence, *GST* π, thrombin, and MCS was obtained from pGSTπ-4T-1 by PCR amplification using the 5′-*Pvu*II and 3′-*Apa*I (5′-CGGCCCGGGCCCAGTCAGTCACGATGCGGC) primers and digestion with *Pvu*II + *Apa*I. The *GST* π expression cassette was inserted between the CMV promoter and the poly(A) signal of the eukaryotic expression plasmid pCMV.neo (Carstens et al., 1995) and digested with *Hin*dIII (repaired with PolIk) and *Apa*I to produce a tandem prokaryotic-eukaryotic expression vector pGFLEX. The cyclin A coding sequence (1299 bp) was PCR amplified from pSVcyclin A (Wang et al., 1990) using 5′-cyclin A-*Bam*HI (5′-GCGCGG-ATCCTTGGGCAACTCTGCGCCGGGGCCTGCG), and 3′-cyclin A-*Apa*I primers (5′-CCAGAGACA-CTAAATCTGTAAGGGCCCGC) and was cloned in-frame into pGFLEX between *Bam*HI (5′) and *Apa*I (3′) sites to generate the GST π-cyclin A fusion protein expression plasmid, pGFLEX-cyclin A.

2.2. Synthesis of GST π fusion proteins in E. coli

E. coli TG1 cells, transformed with the GST π or GST π fusion protein vectors, were grown to early log phase in LB containing Ap (50 μg/ml). Expression of the recombinant proteins was induced by addition of 0.5 mM IPTG, and after 8 h the bacterial cells were harvested by centrifugation (5000 × *g* for 5 min). The cell pellets were washed once with phosphate-buffered saline (PBS, 140 mM NaCl/5 mM KCl/75 mM Na_2HPO_4, pH 7.15) and resuspended in PBS. The cells were lysed by sonication on ice. The sonicated cell extract was centrifuged twice at 10 000 × *g* for 15 min at 2°C and the cleared supernatant was used as the source of GST π or GST π fusion proteins.

2.3. Transient expression in COS-7 cells and preparation of whole cell extract

COS-7 monkey kidney cells were cultured to 50% confluence in Dulbecco's modified Eagle's medium (DMEM) containing 10% fetal calf serum (FCS). The pGFLEX expression constructs were transfected into COS-7 cells using DEAE-dextran (Puchalski and Fahl, 1990) lipofection with DOTAP (*N*-[1-(2,3-dioleoyloxy)propyl]-*N*,*N*,*N*-trimethylammonium methyl sulfate, Waisman Center, UW, Madison, WI, USA) and electroporation. COS-7 cells (5 × 106 per 150 mm dish), grown for 12 h in 150 mm dishes, were exposed to the plasmid DNA (5 μg)/DEAE-dextran (1.4 mg) complex in 10 ml DMEM/20 mM Hepes. The cells were washed once with PBS and treated for 1 min with 10% DMSO in Hepes-buffered PBS (140 mM NaCl/5 mM KCl/75 mM Na_2HPO_4/25 mM Hepes, pH 7.15). The cells were then washed with PBS and supplied with DMEM/10% FCS medium. The lipofection mixture contained 10 μg plasmid DNA and 50 μg DOTAP in 10 ml DMEM medium. The cells were incubated with the plasmid-DOTAP mixture for 6 h and then refed with DMEM/10% FCS medium. Conditions for electroporation were: 5 μg DNA per 5×10^6 cells in 0.4 ml of

DMEM, 250 uF/400 V (Bio-Rad Gene Pulser, Bio-Rad, Hercules, CA, USA). The cells and DNA were incubated on ice for 5 min, then pulsed in a 0.4 cm gap cuvette (Bio-Rad) and transferred to a 150 mm dish containing DMEM/10% FCS. Two days after transfection, the cells were harvested and protein extracts were prepared as previously described (Puchalski and Fahl, 1990). Briefly, the cells were sonicated and the cell lysate was centrifuged twice in a Sorvall SS34 rotor at 10 000 × *g* for 15 min. The cleared cytosolic supernatant protein fraction was retained for further analysis of GST or GST fusion protein.

2.4. Purification of GST π and GST π fusion proteins

For use in enzyme assays, the chimeric GST π was purified on a GSH-Sepharose 4B RediPack column (Pharmacia Biotech, Uppsala, Sweden). The supernatant protein solution was loaded on a GSH-Sepharose 4B column pre-equilibrated with buffer A (10 mM Tris-HCl and 1 mM EDTA pH 7.8). The column was washed three times with 20 ml buffer A + 250 mM NaCl and the chimeric GST π proteins were eluted with buffer A containing 5 mM GSH and dialyzed overnight in 10 mM Tris-HCl/1 mM EDTA (pH 7.2).

For experiments described in Fig. 4, testing the ability of the GST π fusion protein to bind to the GSH affinity matrices, the method of Smith and Corcoran (1994) was followed. Two gram of GSH-Sepharose 4B or *S*-hexyl-GSH-agarose (Sigma, St. Louis, MO, USA) was soaked overnight in 50 ml of buffer A. The hydrated affinity matrix was pelleted by centrifugation at 1000 × *g* for 5 min. After discarding the supernatant the affinity matrix was washed three times each with 50 ml of wash buffer A (Tris-HCl, pH 8.5 + 500 mM NaCl), wash buffer B (Tris-HCl, pH 4.5 + 100 mM sodium acetate + 500 mM NaCl) and buffer A. After equilibration with buffer A overnight, approx. 0.2 g (wet weight) affinity matrix and 100 μg cytosolic protein fraction were incubated in a volume of 1 ml (buffer A) in a 1.5 ml Eppendorf tube for 30 min at 5°C using a rotating platform (Torque-Rotator, Cole-Parmer, Chicago, IL, USA). The suspension was centrifuged at 1000 × *g* for 10 min and the supernatant was discarded. The pellet was washed by resuspending in 1 ml buffer A + 200 mM NaCl and gently shaken on a rotating platform for 10 min. The washing step was repeated twice. The GST bound to the affinity matrix after each washing step was eluted with 200 μl of buffer A + 5 mM GSH or *S*-hexyl-GSH by gentle shaking on a rotating platform for 15 min. After pelleting the affinity matrix by centrifugation at 1000 × *g* for 10 min, the clear supernatant containing the GST was removed using a narrow flat Eppendorf pipet tip. 50 μl aliquots from the GSH/*S*-hexyl-GSH eluted fraction were mixed with an equal volume of 2 × SDS-PAGE loading buffer (125 mM Tris-HCl, pH 6.8, 4% SDS, 10% 2-mercaptoethanol, 20% glycerol and 0.05% bromophenol blue), boiled for 5 min, and resolved on a 12.5% SDS-polyacrylamide gel.

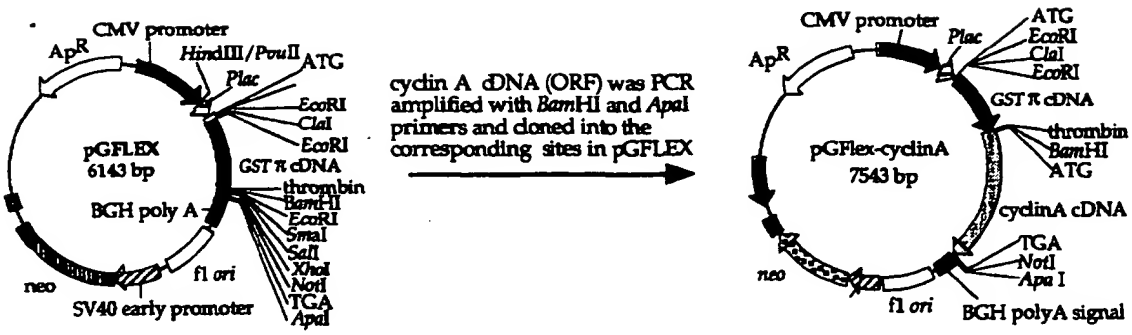
Proteins in the gels were identified by staining with Coomassie blue or by Western blots. Proteins were electrophoretically transferred to nitrocellulose membranes and Western blots were done using the published procedure (Sambrook et al., 1989). GST π or GST π-cyclin A fusion protein and CDK2 were detected using polyclonal rabbit anti-human GST π (Gulick et al., 1992) anti-cyclin A (Oncogene Science, Bethesda, MD, USA) and anti-CDK2 (Upstate Biotech, New York, NY, USA), respectively, as primary Ab (1:5000 dilution) and alkaline phosphatase or horseradish peroxidase conjugated goat anti-rabbit IgG (Jackson Immunoresearch Lab, West Grove, PA, USA; 1:10 000 dilution) as the secondary Ab. The commercially available polyclonal primary Ab though advertised as specific for human cyclin A and CDK2, cross-reacted with some other proteins under our Western blot blocking conditions. Bands were visualized by incubating in BCIP/NBT or luminol reagent (Pierce, Rockford, IL, USA). Protein estimation was done using the BCA reagent (Pierce).

3. Results and discussion

3.1. Synthesis of chimeric GST π in E. coli

Among the bacterial promoters tested in our laboratory (including λP_L, *Ptac* and *pPac*), the *Plac* was the most efficient at expressing a full-length human GST π cDNA (pUC120GSTπ, Manoharan et al., 1992a,b) with a yield of native GST π (210 aa; M_r 23 373) up to 4–8% of the soluble protein in the bacterial cell extract (Fig. 2A, lanes 4, 5). In an attempt to further increase the level of GST π synthesis in *E. coli*, we considered some genetic manipulations which have been utilized to maximize protein production. It has been shown that the presence of specific aa residues at the N terminus of βGal confers stability of this protein in bacterial and mammalian cells (Varshavsky, 1992): Addition of the appropriate aa residues to the N terminus of GST π cDNA, especially a block of 13 N-terminal aa from the engineered βGal leader sequence, increased the level of GST π synthesis to 20–30% of the total soluble proteins (Fig. 2B, lane 5). When the N-terminal aa sequence of the βGal-GST π fusion protein was determined, it revealed the presence of the βGal leader sequence aa in the fusion protein (Fig. 1B). Fusion of the βGal leader sequence provided the GST π expression cassette with the advantages of a strong ATG initiator, an alternate initiator at the third codon, an efficient bacterial RBS, and the stabilizing N-terminal aa Thr which has been shown to confer greater stability to proteins in *E. coli*. Further steps in building the GST π expression cassette

A



B

```
850  RBS
   AC AGGA AACAGCT   ATG ACC ATG ATT ACG AAT TCA TCG ATG AAT TCC GCC ACC ATG CCC
                     Met Thr Met Ile Thr Asn Ser Ser Met Asn Ser Ala Thr Met Pro
                    1   2   3  4   5   6   7   8   9  10  11  12  13
                   ------------ βGal leader sequence -------------- ---GSTπ---
```

C

```
AAA CAG CTG GTT CCG CGT GGA TCC CCG GAA TTC CCG GGT CGA CTC GAG CGG CCG CAT CGT GACTGACTGGGCCCTATC
Lys Gln Leu Val Pro Arg Gly Ser Pro Glu Phe Pro Gly Arg Leu Glu Arg Pro His Arg Asp***polyA signal
-GSTπ-      thrombin       BamHI        EcoRI  SmaI   SalI  XhoI    NotI                    ApaI
```

Fig. 1. (A) Schematic representation of the of GST π fusion expression vectors. Construction details are available on request. (B) A portion of the *lac* operator and transcription initiation region showing the nt and aa sequences of the βGal leader and human *GST π*. (C) The nt sequence of pGFLEX at the C terminus of *GST π* showing the thrombin cleavage and MCS.

of pGFLEX included adding a 6 aa thrombin cleavage sequence and a MCS to the C terminus. These manipulations added an additional 13 aa (Fig. 1C) to the protein but did not affect the amount of GST π (hereafter referred to as chimeric GST π) produced (Fig. 2B, lane 6). The amount of chimeric GST π produced was at least as high as that produced by the pGEX-4T-1 vector which encodes a schistosomal GST (Smith and Johnson, 1988) (Fig. 2B, lane 4). Expression of human GST π from the engineered vectors was not lethal to *E. coli* cells as judged by a normal growth rate of bacteria during the course of enzyme induction with IPTG.

Subcellular fractionation of protein extracts from pGFLEX-transformed *E. coli* cells revealed that 80–90% of the chimeric GST π was in the soluble fraction. This indicated that the chimeric GST π was not incorporated into inclusion bodies. The size of the chimeric GST π estimated by SDS-PAGE is 28 kDa (Fig. 2B, lane 6) matching the size of the chimeric GST π (242 aa; M_r 27 613) predicted from the coding capacity of the βGal leader/GST π/T/MCS expression cassette.

Comparison of the biochemical properties of the bacterially produced chimeric GST π to the wt GST π purified from human placenta indicated that the chimeric GST π was as active as the wt enzyme. The specific activity of the chimeric GST π for the 1-chloro-2-4-dinitrobenzene (CDNB)-GSH conjugation reaction was slightly greater (1.2-fold) than that of the human placental GST π (15.2 units/μg protein). Additionally, the chimeric GST π bound efficiently to both GSH and *S*-hexyl-GSH affinity matrices, and the chimeric GST π retained in the affinity matrix could be recovered by elution with GSH or *S*-hexyl-GSH.

3.2. Synthesis of chimeric GST π in mammalian cells

In addition to the sequences necessary for efficient bacterial expression, pGFLEX includes the CMV promoter and a polyadenylation signal flanking the GST π cassette for expression in mammalian cells (Fig. 1A). The vector also contains the SV40 *ori* and early promoter fused to the Nm^R gene. These two features facilitate the rapid amplification of pGFLEX in the large T-antigen expressing COS-7 cells and the selection for transfected cells using G418. To test the ability of pGFLEX to support expression of chimeric GST π in mammalian cells, the plasmid was transfected into COS-7 cells by DEAE-dextran and lipofection protocols. Following transfection (72 h), high levels of chimeric GST π were observed. Efficient transfection of COS-7 cells was achieved with the DEAE-dextran method, and the high-level expression (6–8% of total protein extract) of the

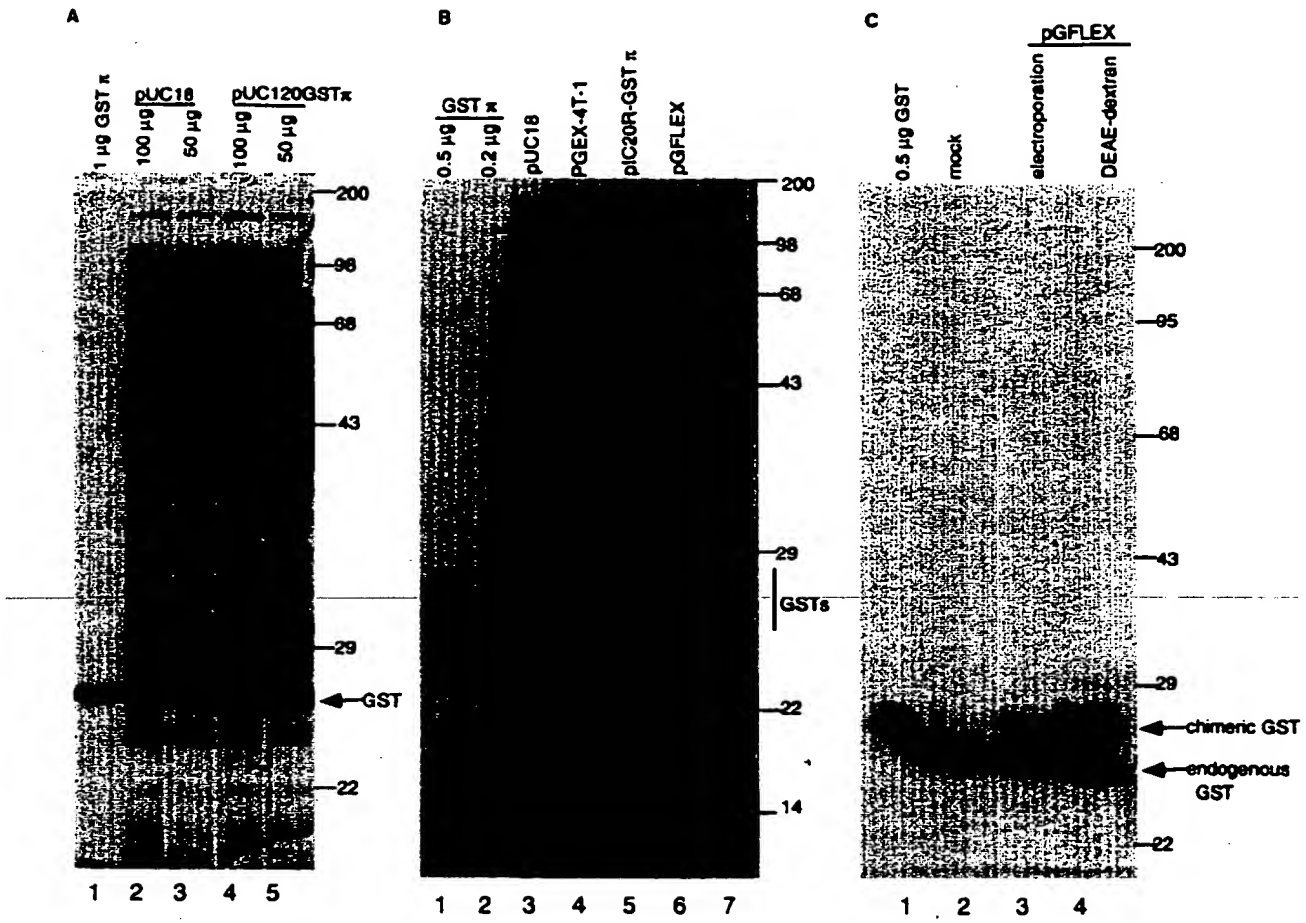


Fig. 2. (A, B) Synthesis of GST π in *E. coli*. *E. coli* cells containing expression vectors were induced with IPTG and protein extracts were prepared as described in Section 2. Each lane was loaded with the amount of protein indicated in the figure, the proteins were separated by SDS-PAGE (12.5%) and stained with Coomassie blue. Purified human GST π was used as a standard for comparison. Positions of molecular weight markers are shown on the right. (C) Transient synthesis of recombinant GST π in COS-7 cells. Cells were transfected with pGFLEX by electroporation or DEAE-dextran. After 48 h, soluble cytosolic protein was prepared from mock transfected (lane 2), lipofection (lane 3) and DEAE-dextran transfected (lane 4) cells. Each lane was loaded with 50 μg of protein. The amount of chimeric GST π produced in the cells was analyzed by Western blot using anti-GST π and alkaline phosphatase conjugated anti-rabbit Ab. 0.5 μg purified human GST π was used (lane 1) as a marker.

28-kDa chimeric GST π (Fig. 2C, lane 4) could be distinguished from the faster-migrating (23 kDa) endogenous GST π (Fig. 2C, lane 2) in the monkey kidney cells. No immunoreactive, proteolytic degradation fragments were observed. The presence of the 104-bp *Plac* sequence between the CMV promoter and the βGal leader/GST π cassette did not prevent the efficient synthesis of chimeric GST π in mammalian cells. pGFLEX also enabled production of chimeric GST π in 3T3 cells but at a lower level than in COS-7 cells (data not shown).

3.3. Synthesis of target fusion proteins

The pGFLEX system was then tested for its ability to produce fusion proteins in *E. coli* and mammalian cells, and the target fusion protein produced was further evaluated for the retention of its biological activities.

Cyclin A is a protein synthesized in all eukaryotic cells and is known to be the integrator of the signal transduction pathways that drive the arrest or activation of cell division. During these processes, cyclin A has been shown to interact with multiple proteins including cyclin-dependent kinase 2 (CDK2; apparent molecular mass 33 kDa), E2F, p21, proliferating cell nuclear antigen (PCNA), p107, and the adenovirus-transforming protein E1A (Peeper et al., 1994). In vitro experiments have shown that at the onset of cell division there is an ordered activation of CDK2 and the activity of the kinase is controlled in part through complex formation with a regulatory subunit, cyclin A (Elledge et al., 1992). There is a wealth of knowledge concerning protein-

protein interactions between cyclin A and other cellular proteins, and Ab to cyclin A (apparent M_r 48.5 kDa; Fig. 3C, lane 1) and its protein partners are commercially available. Because of this, we chose to use cyclin A interactions as a paradigm for testing the ability of the pGFLEX system to enable production and rapid purification of a biologically functional GST π-cyclin A fusion protein from mammalian cells.

Despite the fact that cyclins are inherently unstable proteins (Hochastrasser, 1995) construction of pGFLEX-cyclin A (Fig. 1A) and its expression in *E. coli* (Fig. 3A, lane 4) or COS-7 (Fig. 3B, lanes 2, 3; Fig. 3C, lane 3) cells resulted in readily detectable amounts of the 76-kDa GST π-cyclin A fusion protein. However, low levels of partial proteolysis of the fusion protein were noticed in few batches of protein extracts. The DEAE-dextran transfection proved to be more efficient than either electroporation (Fig. 3B, lane 2) or lipofection (Fig. 2C, lane 3) transfection of pGFLEX-cyclin A into COS-7 cells. Cleavage of the fusion protein with thrombin released cyclin A (48.5 kDa) from chimeric GST π (Fig. 3A, lane 5). The GSTπ-cyclin A protein retained strong reactivity with GST π and cyclin A Ab (Fig. 3B and C).

Since our goal was to rapidly purify the fusion protein, the GSH binding affinity of the GST π-cyclin A fusion protein was examined. Different isozymes of GST have been shown to bind GSH and *S*-hexyl-GSH affinity matrices with different affinities (Mannervick and Danielson, 1988). We therefore tested both affinity matrices for their abilities to bind the GST π-cyclin A fusion protein. The GST π-cyclin A fusion protein bound to both GSH (Fig. 4A, lanes 9–12) and *S*-hexyl-GSH (Fig. 4B, lanes 9–12) affinity matrices equally well but did so less efficiently than either native GST π (Fig. 4A and B, lanes 1–4) or chimeric GST π (Fig. 4,

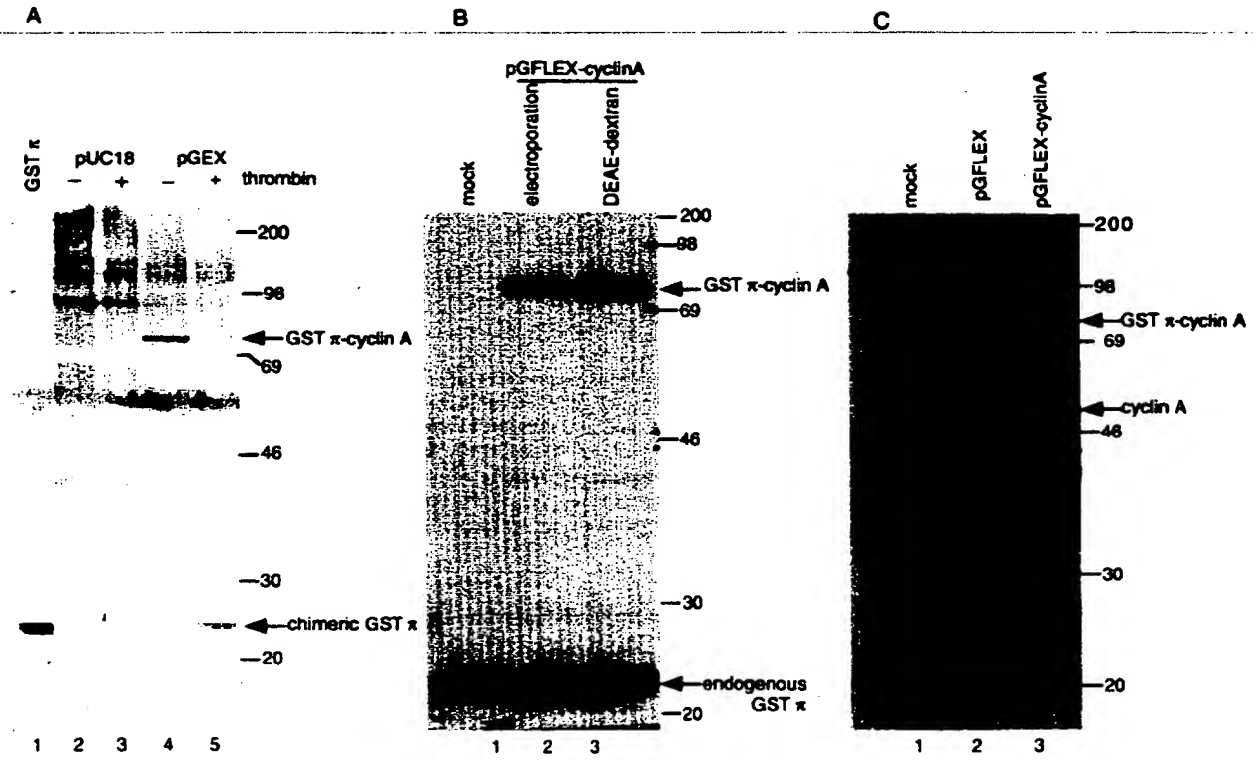


Fig. 3. (A) Synthesis of GST π-cyclin A fusion protein in *E. coli*. *E. coli* cultures (2 ml) containing pUC18 (lanes 2 and 3) or pGFLEX-cyclin A (lanes 4 and 5) plasmids were induced with IPTG. After 8 h, 0.5 ml of culture was removed, and the cells were pelleted in an Eppendorf centrifuge. The cell pellets were resuspended in 400 μl PBS and lysed by sonication. After centrifuging at 10 000 × *g* for 10 min, two aliquots (20 μl each) of the cleared supernatants were taken and incubated for 10 min at 37°C either in the presence (lanes 3 and 5) or absence (lanes 2 and 4) of thrombin (20 units; 10 units/μl, Sigma). The samples were mixed with an equal volume of 2 × SDS-PAGE loading buffer and separated on a 10% SDS-PAGE gel. The GST π and GST π-cyclin A fusion proteins were visualized by Western blot using a polyclonal rabbit anti-human GST π Ab and alkaline phosphatase conjugated goat anti-rabbit Ab as primary and second Ab, respectively. Purified human GST π (0.1 μg) was used as a marker. (B, C) Synthesis of GST π-cyclin A fusion protein in COS-7 cells. COS-7 cells were transfected with pGFLEX (C, lane 2) or pGFLEX-cyclin A using DEAE-dextran (B, lane 3) or electroporation (B, lane 2) or mock transfection (B, lane 1; C, lane 1). Cells were harvested 48 h post transfection. Aliquots of cytosolic protein (50 μg) were analyzed for the level of GST π-cyclin A fusion protein by Western blot using (B) a polyclonal rabbit anti-GST π and a goat anti-rabbit Ab conjugated to alkaline-phosphatase, or (C) a polyclonal rabbit anti-human cyclin A Ab and goat anti-rabbit Ab conjugated with horseradish peroxidase.

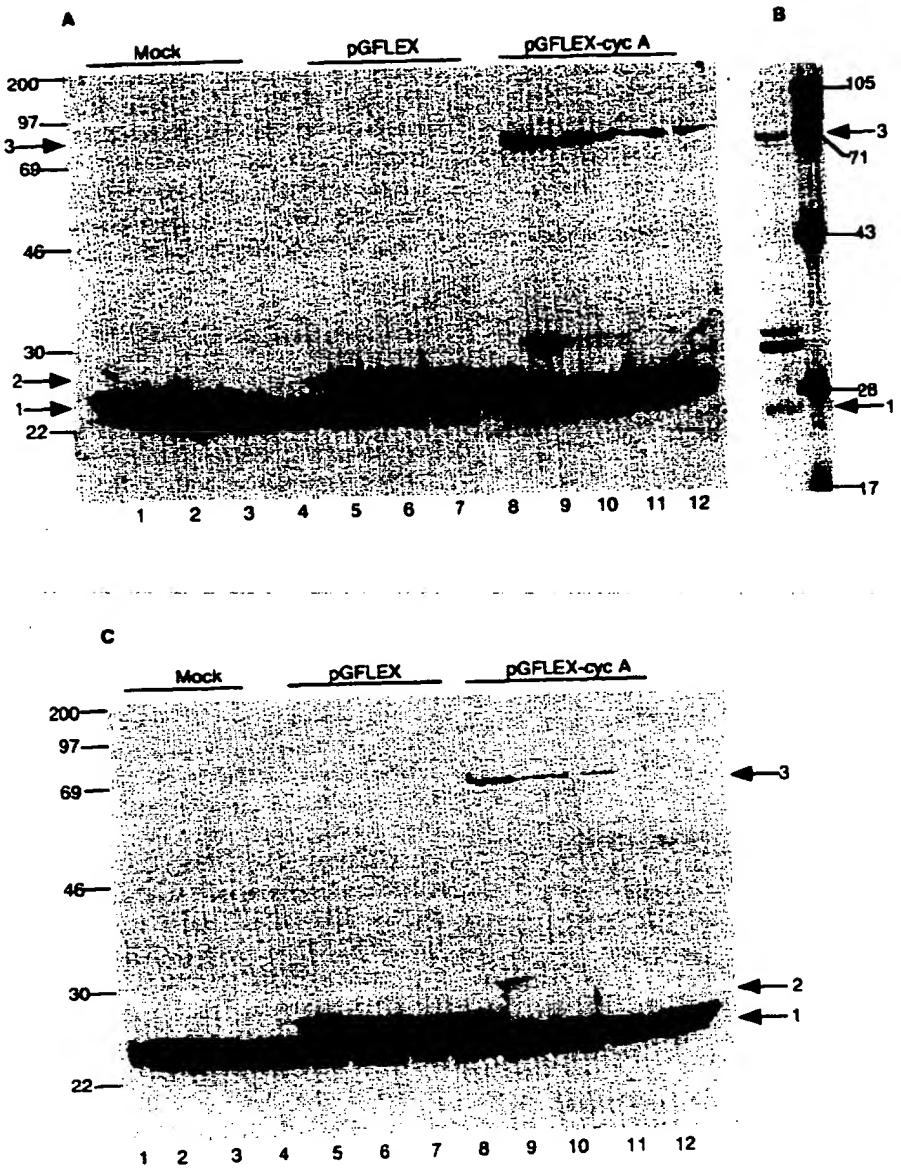


Fig. 4. Capacity of chimeric GST π and GST π-cyclin A fusion protein to bind GSH-Sepharose 4B (A, Western blot; B, Coomassie-stained gel of purified GST π-cyclin A fusion protein from COS-7 cells) and *S*-hexyl-GSH-agarose (C). Cytosolic protein extracts (100 μg each) from mock (lanes 1–4), pGFLEX (lanes 5–8) or pGFLEX-cyclin A (lanes 9–12) transfected COS-7 cells were incubated with affinity matrix as described in Section 2. GST π and GST π-cyclin A fusion proteins were eluted directly with GSH/*S*-hexyl-GSH (lanes 1, 5, 9) or after washing once (lanes 2, 6, 10), twice (lanes 3, 7, 11) or thrice (lanes 4, 8, 12) with buffer A+250 mM NaCl. To obtain maximum sensitivity for detection of GST π (arrow 1), chimeric GST π (arrow 2) and GST π-cyclin A fusion protein (arrow 3) eluted from the affinity matrix, we used rabbit anti-GST π and horseradish peroxidase conjugated goat anti-rabbit IgG Ab for the Western blots.

lanes 5–8) when tested under more stringent conditions. Nonetheless, the GST π-cyclin A fusion protein's ability to bind to a GSH affinity matrix is sufficient to allow purification of the protein from bacterial and mammalian cells.

We next determined whether the cyclin A component of the GST π-cyclin A fusion protein retained its biochemical properties. As one measure of biological activity, the association of cyclin A with its CDK2 partner (Elledge et al., 1992) was examined. COS-7 cells were transfected with pGFLEX-cyclin A and the cell-free protein extract was passed over a GSH-Sepharose affinity column. After extensive washing of the affinity column with buffer A containing 100 mM NaCl (low stringency condition), the proteins retained by the GSH-Sepharose were examined by immunoblotting.

Immunoblot analysis using an anti-CDK2 Ab revealed co-retention of CDK2 (33 kDa) by the GST π-cyclin A fusion protein (Fig. 5, lane 6). No detectable CDK2 was observed in either untransfected COS-7 cells or the COS-7 cells transfected with pGFLEX (Fig. 5, lanes 4, 5). These results indicate that the GST π-cyclin A fusion protein synthesized from pGFLEX-cyclin A remains functional in mammalian cells, and that the protein-protein interactions mediating the association of cyclin A and CDK2 previously demonstrated by in vitro studies can also be detected in vivo using the fusion protein encoded by our vector. A pGFLEX-luc construct enabled expression of a GST π-Luc fusion protein, and the GST π-Luc protein was also easily purified using a GSH affinity matrix (data not shown).

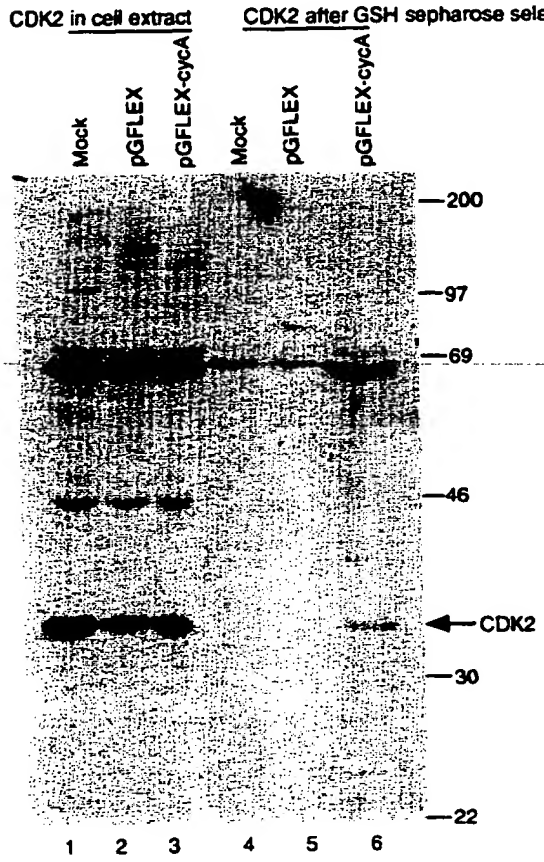


Fig. 5. In vivo association of GST π-cyclin A fusion protein with CDK2. pGFLEX-cyclin A was transfected using DEAE-dextran into COS-7 cells. Cells were harvested 48 h post transfection, washed with PBS and lysed with lysis buffer (Tris-HCl, pH 7.2, 1 mM EDTA, 0.1% Nonidet P-40) and the nuclei were disrupted by extruding the cell lysate through a 22 G needle six times. Lysates were centrifuged at 1000 × *g* for 10 min, and the clear cytosolic supernatant fraction (100 μg protein) was mixed with 0.2 mg GSH-Sepharose 4B for 30 min in a 1.5 ml Eppendorf tube in a total volume of 1 ml buffer A on a rotating platform at 5°C. The GST π-cyclin A fusion protein and the proteins complexed with the GST π-cyclin A fusion protein were washed three times with 1 ml wash buffer (10 mM Tris-HCl, pH 7.8, 100 mM NaCl). Proteins complexed with GST π-cyclin A fusion protein that was bound to the pelleted GSH affinity matrix were eluted by boiling the pellet in 100 μl of 1 × SDS-PAGE loading buffer. The affinity matrix was pelleted by centrifuging at 10 000 × *g* for 15 min. Supernatant (50 μl) was electrophoresed on a 12.5% SDS-acrylamide gel, and CDK2 was detected by Western blot using a rabbit anti-CDK2 and horseradish peroxidase conjugated goat anti-rabbit IgG as primary and second Ab, respectively. The amount of CDK2 present in the protein extracts (100 μg each; lanes 1–3) and retained by affinity matrix (lanes 4–6) are shown by the arrow.

4. Conclusions

The pGFLEX is a unique vector enabling synthesis and purification of target proteins ranging from a single polypeptide to one or more components of a multi-protein complex in both *E. coli* and mammalian cells.

Many of the mammalian GST expression vectors previously constructed in our laboratory, utilizing the SV40 immediate-early, MMTV-*LTR*, RSV-*LTR*, α-collagen and HSV-*tk* promoters, enabled high-level expression of native GST π in mammalian cells (Gulick and Fahl, 1995). In contrast, maximizing high-level production of native GST π in *E. coli*, to a level comparable to that of pGEX-4T-1, required several genetic manipulations. A eukaryotic schistosomal GST fusion vector containing SV40 early enhancer/promoter with the second intron of the rabbit β-globin gene was used for the study of ATFa interaction with Jun and Fos (Chatton et al., 1995). Our early attempts to develop a bacterial-mammalian shuttle expression vector utilizing the schistosomal GST from pGEX-4T-1 were not successful. Of the several pGEX-4T-1-based eukaryotic expression vectors constructed in our laboratory, only the vector containing the schistosomal GST cDNA fused to the CMV promoter produced barely detectable amounts of schistosomal GST and extensive degradation of schistosomal GST was observed in the mammalian cell background.

In addition to the benefits of producing high levels of a target protein, the pGFLEX vector may find application in in vivo studies that aim to retrieve and identify new proteins of multicomponent complexes which interact with the target protein of interest. pGFLEX provides researchers with a tool that can be easily used for in vivo and in vitro studies to define the domains of proteins that mediate protein-protein interactions and dimerization. The vector also offers a means for the rapid purification of large quantities of a protein that associates with the target protein. The vector possesses the f1 replication origin so that single strand DNA can be produced allowing mutagenesis to be performed easily using conventional methods.

We have introduced a unique *ClaI* cloning site immediately upstream from the *GST* π cDNA to facilitate the expression of target proteins at the GST N terminus

(Fig. 1A). We have successfully produced active cyclin A protein at the N terminus of GST π by inserting the full-length *cyclin* A cDNA into the *Cla*I site in-frame with the *GST* π cDNA sequence. The *Cla*I site can also be used to incorporate specific targeting signal sequences that direct the fusion proteins to different subcellular organelles. We have inserted the nuclear localization signal (NLS-1, a block of 28 aa) of the glucocorticoid receptor (Rusconi and Yamomoto, 1987) into the *Cla*I and observed NLS-GST localization in the nuclei of COS-7 cells (Kramer et al., 1995). Attempts are underway to produce GST π on the mammalian cell surface as a CD4 fusion protein. Extracellular expression of the GST π-CD4 fusion protein might provide a unique handle for identifying and purifying a population of transfected cells that co-express the GST π-CD4 fusion protein and a transfected gene of interest.

Acknowledgement

The authors thank Greg Kujoth for his advice in the preparation of this manuscript. This research was supported by grants PO1-CA22484 and P30-CA07175 from the National Institutes of Health.

References

Carstens, C-P., Gallo, J.C., Maher, V.M., McCormick, J., Fahl, W.E., 1995. A system utilizing Epstein-Barr virus based expression vectors for the functional cloning of human fibroblast growth regulators. Gene 164, 195–202.

Chatton, B., Bahr, A., Acker, J., Kedinger, C., 1995. Eukaryotic GST fusion vector for the study of protein-protein associations in vivo: application to interaction of ATFa with Jun and Fos. Biotechniques 18, 142–145.

Datur, R.V., Cartwright, T., Rosen, G-R., 1993. Process economics of animal cell and bacterial fermentations: a case study analysis of tissue plasmogen activator. Bio/Technology 11, 349–357.

de Wet, J.R., Wood, K.V., DeLuca, M., Helinski, D.R., Subramani, S., 1987. Firefly luciferase gene: structure and expression in mammalian cells. Mol. Cell. Biol. 7, 725–737.

Elledge, S.J., Richman, R., Hall, F.L., Williams, R.T., Lodgson, N., Harper, W., 1992. CDK2 encodes a 33-kDa cyclin A-associated protein kinase and is expressed before CDC2 in the cell cycle. Proc. Natl. Acad. Sci. USA 89, 2907–2911.

Gold, L., 1990. Expression of heterologous proteins in *Escherichia coli*. Methods Enzymol. 185, 11–14.

Gulick, A.M., Fahl, W.E., 1995. Mammalian glutathione *S*-transferase: regulation of an enzyme system to achieve chemotherapeutic efficacy. Pharmacol. Ther. 66, 237–257.

Gulick, A.M., Goihl, A.L., Fahl, W.E., 1992. Structural studies on human glutathione *S*-transferase π. Family of native-specific monoclonal antibodies used to block catalysis. J. Biol. Chem. 267, 18946–18952.

Hewick, R.M., Hunkapiller, M.W., Hood, L.E., Dreyer, W.J., 1981. A gas-liquid phase peptide and protein sequenator. J. Biol. Chem. 256, 7990–7997.

Hochastrasser, M., 1995. Ubiquitin, proteasomes, and the regulation of intracellular protein degradation. Curr. Opin. Cell Biol. 7, 215–223.

Kano, T., Sakai, M., Muramatsu, M., 1987. Structure and expression of a human class π glutathione *S*-transferase messenger RNA. Cancer Res. 47, 5626–5630.

Kramer, K.S., Manoharan, T.H., Schumacher, C. and Fahl, W.E., 1995. Nuclear localization of recombinant glutathione *S*-transferase to improve drug detoxification. Gene Therapy and Molecular Medicine. Key Stone Symposium, Key Stone, CO, June 10–15, p. 272.

Mannervick, B., Danielson, U.H., 1988. Glutathione transferases – structure and catalytic activity. CRC Crit. Rev. Biochem. 23, 283–337.

Manoharan, T.H., Gulick, A.M., Reinemer, P., Dirr, H.W., Huber, R., Fahl, W.E., 1992. Mutational substitution of residues implicated by crystal structure in binding the substrate glutathione to human glutathione *S*-transferase π. J. Mol. Biol. 226, 319–322.

Manoharan, T.H., Gulick, A.M., Puchalski, R.B., Servis, A.L., Fahl, W.E., 1992. Structural studies on human glutathione *S*-transferase π. Substitution mutations to determine amino acids necessary for binding glutathione. J. Biol. Chem. 267, 18940–18945.

Marsh, J.H., Erfle, M., Wykes, E., 1984. The pIC plasmid and phage vectors with versatile cloning sites for recombinant selection by insertional inactivation. Gene 32, 481–485.

Maurizi, M.R., 1992. Proteases and protein degradation in *Escherichia coli*. Experientia 48, 178–201.

Mitraki, A., Haase-Pettingell, C., Sturtevant, J., King, J., 1991. Global suppression of protein folding defects and inclusion body formation. Science 253, 54–58.

Olins, P.O., Lee, S.C., 1993. Recent advances in heterologous gene expression in *Escherichia coli*. Curr. Opin. Biotechnol. 4, 520–525.

Peeper, D.S., Van der Eb, E.J., Zantema, A., 1994. The G1/S cell-cycle checkpoints in eukaryotic cells. Biochim. Biophys. Acta 1198, 215–230.

Puchalski, R.B., Fahl, W.E., 1990. Expression of recombinant glutathione *S*-transferase π, Ya or Yb1 confers resistance to alkylating agents. Proc. Natl. Acad. Sci. USA 87, 2443–2447.

Rusconi, S., Yamomoto, Y.R., 1987. Functional dissection of the hormone and DNA binding activities of glucocorticoid receptor. EMBO J. 6, 1309–1315.

Sambrook, J., Fritsch, E.F. and Maniatis, T., 1989. Molecular Cloning. A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Smith, D.B. and Corcoran, L.M., 1994. Expression and purification of glutathione *S*-transferase fusion protein. In: Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A. and Struhl, K. (Eds.), Current Protocols in Molecular Biology. John Wily and Sons, New York, NY, pp. 16.7.2–16.7.7.

Smith, D.B., Johnson, K.S., 1988. Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione *S*-transferase. Gene 67, 31–40.

Varshavsky, A., 1992. The N-end rule. Cell 69, 725–735.

Wang, C.X., Henglein, B., Brechot, C., 1990. Hepatitis B virus integration in a cyclin A gene in a hepatocellular carcinoma. Nature 343, 555–557.